

# Analisis Performa Support Vector Machine untuk Klasifikasi Risiko Kredit Nasabah pada Perbankan Daerah

Rahmat, Asep Sapaatullah<sup>2</sup>, Mochammad Darip<sup>3,\*</sup>

<sup>1</sup> Fakultas Informatika, Program Studi Manajemen Informatika, Akademi Manajemen Informatika Dan Komputer, Kota Serang - Banten, Indonesia

<sup>2</sup> Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Bina Bangsa, Kota Serang - Banten, Indonesia

<sup>3</sup> Fakultas Ilmu Komputer, Program Studi Ilmu Komputer, Universitas Bina Bangsa, Kota Serang - Banten, Indonesia

Email: <sup>1</sup>[rahmat042@gmail.com](mailto:rahmat042@gmail.com), <sup>2</sup>[asepsapaatullah.binabangsa@gmail.com](mailto:asepsapaatullah.binabangsa@gmail.com),

(\* : [darif.uniba@gmail.com](mailto:darif.uniba@gmail.com) )

**Abstrak-** Penilaian risiko kredit merupakan komponen krusial dalam sistem perbankan karena berhubungan langsung dengan kemampuan lembaga keuangan dalam mengelola potensi kerugian akibat kredit bermasalah. Permasalahan yang sering dihadapi bank adalah kesulitan dalam mengklasifikasikan tingkat risiko kredit nasabah secara akurat, terutama ketika data yang dianalisis bersifat kompleks, nonlinier, dan memiliki variabel yang saling berinteraksi. Metode konvensional seperti analisis regresi sering kali kurang mampu menangkap pola-pola tersembunyi dalam data tersebut. Oleh karena itu, penelitian ini bertujuan untuk menerapkan algoritma Support Vector Machine (SVM) sebagai solusi untuk mengklasifikasikan tingkat risiko kredit nasabah bank berdasarkan atribut seperti pendapatan, jumlah pinjaman, lama bekerja, riwayat pembayaran, rasio utang terhadap pendapatan, serta status kepemilikan aset. Proses penelitian diawali dengan pengumpulan dan pra-pemrosesan data, termasuk pembersihan data dan normalisasi untuk memastikan distribusi nilai yang seragam. Selanjutnya dilakukan pembagian data menjadi data latih dan data uji dengan proporsi tertentu, kemudian diterapkan model SVM menggunakan beberapa jenis kernel seperti linear, polynomial, dan radial basis function (RBF) untuk menentukan kernel dengan kinerja terbaik. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score guna mengukur tingkat performa klasifikasi. Hasil pengujian menunjukkan bahwa model SVM dengan kernel RBF memberikan hasil terbaik dengan tingkat akurasi mencapai lebih dari 90%, serta mampu meminimalkan kesalahan klasifikasi pada kategori risiko tinggi. Kesimpulannya, penerapan algoritma SVM terbukti efektif dalam mengklasifikasikan tingkat risiko kredit nasabah dengan tingkat akurasi dan stabilitas yang tinggi, sehingga dapat dijadikan sebagai alat bantu yang andal bagi pihak bank dalam proses analisis kelayakan kredit serta pengambilan keputusan strategis yang lebih tepat dan berbasis data.

**Kata Kunci:** Support Vector Machine, Risiko Kredit, Klasifikasi Nasabah, Analisis Data, Perbankan.

**Abstract-** Credit risk assessment is a crucial component of the banking system because it directly relates to a financial institution's ability to manage potential losses due to non-performing loans. Banks often face difficulties in accurately classifying customer credit risk levels, especially when the data being analyzed is complex, nonlinear, and contains interacting variables. Conventional methods such as regression analysis often fail to capture hidden patterns in such data. Therefore, this study aims to apply the Support Vector Machine (SVM) algorithm as a solution to classify bank customers' credit risk levels based on attributes such as income, loan amount, length of employment, payment history, debt-to-income ratio, and asset ownership status. The research process begins with data collection and pre-processing, including data cleaning and normalization to ensure a uniform distribution of values. The data is then divided into training and test data with specific proportions. An SVM model is then applied using several kernel types, such as linear, polynomial, and radial basis function (RBF), to determine the best-performing kernel. Model evaluation is performed using accuracy, precision, recall, and F1-score metrics to measure classification performance. Test results show that the SVM model with the RBF kernel provided the best results, achieving an accuracy rate of over 90% and minimizing classification errors in the high-risk category. In conclusion, the application of the SVM algorithm has proven effective in classifying customer credit risk levels with high accuracy and stability, making it a reliable tool for banks in the creditworthiness analysis process and more accurate, data-driven strategic decision-making.

**Keywords:** : Support Vector Machine, Credit Risk, Customer Classification, Data Analysis, Banking

## 1. PENDAHULUAN

Bank daerah memiliki karakteristik operasional yang berbeda dibandingkan bank nasional, terutama dalam struktur nasabah dan pola penyaluran kredit, karena sebagian besar nasabah berasal dari sektor usaha mikro dan kecil, pegawai daerah, serta masyarakat dengan variasi tingkat pendapatan yang berbeda [1]. Kondisi tersebut menyebabkan profil risiko kredit menjadi heterogen dan menuntut proses evaluasi yang cermat. Pada bank regional seperti Bank Serang, ketepatan dalam mengklasifikasikan tingkat risiko kredit merupakan faktor penting untuk menjaga kualitas portofolio pinjaman dan mengendalikan rasio kredit bermasalah [2].

Penilaian risiko kredit pada dasarnya bertujuan untuk memperkirakan kemungkinan terjadinya gagal bayar berdasarkan karakteristik finansial dan riwayat pembayaran nasabah [3]. Variabel seperti pendapatan, lama bekerja, jumlah pinjaman, rasio utang terhadap pendapatan, serta riwayat keterlambatan pembayaran sering digunakan sebagai indikator utama [4]. Namun, hubungan antarvariabel tersebut tidak selalu bersifat linier. Sebagai contoh, peningkatan pendapatan tidak selalu berbanding lurus dengan penurunan risiko apabila disertai dengan kenaikan beban utang yang signifikan. Kompleksitas semacam ini menuntut pendekatan analisis yang mampu menangkap pola hubungan yang tidak sederhana.

Pada praktiknya, sebagian proses penilaian kredit di perbankan daerah masih memanfaatkan metode statistik konvensional atau sistem skor berbasis aturan tertentu [5]. Pendekatan tersebut relatif mudah diterapkan, tetapi memiliki

keterbatasan dalam memodelkan interaksi nonlinier antarvariabel [6]. Ketika jumlah atribut semakin banyak dan karakteristik data semakin beragam, model linier cenderung menghasilkan batas keputusan yang kurang representative [7]. Hal ini berpotensi menurunkan ketepatan klasifikasi, terutama pada kategori risiko menengah dan tinggi yang memiliki implikasi langsung terhadap kebijakan pemberian kredit.

Perkembangan machine learning menawarkan pendekatan alternatif dalam permasalahan klasifikasi risiko kredit. Salah satu algoritma yang banyak digunakan dalam konteks klasifikasi adalah Support Vector Machine (SVM) [8]. Metode ini bekerja dengan membentuk hyperplane yang memaksimalkan margin antar kelas, sehingga diperoleh batas keputusan yang optimal [9]. Melalui fungsi kernel, SVM dapat memproyeksikan data ke ruang berdimensi lebih tinggi untuk menangani pola yang tidak dapat dipisahkan secara linier [10]. Karakteristik tersebut menjadikan SVM relevan untuk data risiko kredit yang memiliki struktur kompleks.

Beberapa penelitian sebelumnya menunjukkan bahwa SVM mampu memberikan tingkat akurasi yang kompetitif dalam studi credit scoring [11]. Penerapan SVM pada data pinjaman konsumen dan peringkat kredit perusahaan menunjukkan performa yang baik dibandingkan pendekatan tradisional. Namun demikian, sebagian besar penelitian tersebut menggunakan dataset berskala besar atau berasal dari institusi keuangan dengan karakteristik berbeda dari bank daerah. Perbedaan struktur ekonomi lokal, skala pinjaman, serta profil demografis nasabah dapat memengaruhi pola risiko yang terbentuk. Oleh karena itu, hasil penelitian pada konteks yang berbeda tidak dapat secara langsung digeneralisasi ke perbankan daerah. Selain itu, penelitian sebelumnya umumnya belum secara spesifik mengevaluasi perbandingan performa berbagai jenis kernel SVM serta proses optimasi parameter pada data nasabah bank daerah yang memiliki karakteristik lebih heterogen dan berskala terbatas, sehingga penelitian ini menekankan pada analisis komparatif kernel dan penyesuaian parameter untuk memperoleh model yang paling sesuai dengan kondisi perbankan lokal.

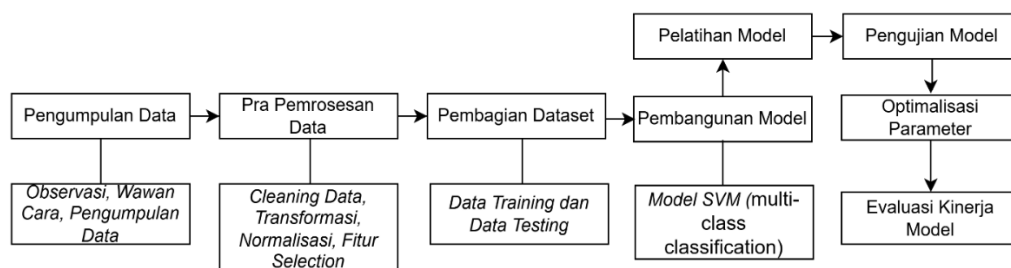
Selain itu, performa SVM sangat dipengaruhi oleh pemilihan jenis kernel dan parameter model [12]. Kernel linear, polynomial, dan radial basis function (RBF) memiliki karakteristik pemetaan yang berbeda terhadap data [13]. Pada beberapa studi, kernel RBF sering menunjukkan kinerja lebih baik untuk data nonlinier, namun efektivitasnya tetap bergantung pada distribusi data yang digunakan. Kajian komparatif yang mengevaluasi performa masing-masing kernel pada data nasabah bank daerah masih relatif terbatas. Padahal, pemilihan kernel yang tidak tepat dapat menghasilkan model dengan tingkat kesalahan klasifikasi yang tinggi.

Berdasarkan kondisi tersebut, penelitian ini difokuskan pada analisis performa Support Vector Machine dalam mengklasifikasikan tingkat risiko kredit nasabah menggunakan data perbankan daerah. Penelitian ini tidak hanya menerapkan SVM sebagai model klasifikasi, tetapi juga membandingkan kinerja beberapa jenis kernel serta melakukan penyesuaian parameter untuk memperoleh konfigurasi yang paling sesuai dengan karakteristik data. Dengan menggunakan data riil nasabah bank daerah, penelitian ini memberikan gambaran empiris mengenai efektivitas SVM dalam nasabah bank lokal.

Tujuan penelitian ini adalah menganalisis kemampuan SVM dalam mengklasifikasikan nasabah ke dalam kategori risiko rendah, sedang, dan tinggi secara terukur. Secara rinci, penelitian ini bertujuan untuk: (1) mengidentifikasi atribut yang relevan dalam penilaian risiko kredit nasabah bank daerah, (2) mengevaluasi performa beberapa jenis kernel SVM melalui proses pelatihan dan pengujian model, serta (3) mengukur kinerja klasifikasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian diharapkan dapat menjadi dasar pertimbangan dalam pengembangan sistem pendukung keputusan kredit yang lebih sesuai dengan karakteristik perbankan daerah.

## 2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi berbasis Support Vector Machine (SVM) untuk menganalisis performa model dalam mengklasifikasikan tingkat risiko kredit nasabah pada perbankan daerah [14]. Fokus penelitian ini adalah membandingkan kinerja beberapa jenis kernel SVM dalam memisahkan kategori risiko kredit berdasarkan karakteristik finansial nasabah. Tahapan penelitian dilakukan secara sistematis meliputi pengumpulan data, pra-pemrosesan, pembagian dataset, pembangunan model SVM dengan variasi kernel, optimasi parameter, serta evaluasi kinerja model [15].



**Gambar 1** Tahapan Penelitian

## 2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data nasabah bank daerah yang memuat informasi karakteristik finansial dan histori kredit. Variabel yang digunakan sebagai atribut prediktor meliputi pendapatan bulanan, jumlah pinjaman, lama bekerja, rasio utang terhadap pendapatan, riwayat pembayaran, dan status kepemilikan asset. Setiap data nasabah telah dikategorikan ke dalam tiga tingkat risiko kredit, yaitu risiko rendah, risiko sedang, dan risiko tinggi. Kategori tersebut ditentukan berdasarkan kebijakan internal bank yang mempertimbangkan riwayat pembayaran dan tingkat kolektibilitas kredit. Sebelum dilakukan pemodelan, data dianalisis secara deskriptif untuk memastikan kelengkapan atribut serta kesesuaian distribusi kelas risiko.

## 2.2 Pra Pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk meningkatkan kualitas data dan memastikan seluruh atribut siap digunakan dalam proses klasifikasi [16]. Langkah-langkah yang dilakukan meliputi:

1. Pembersihan Data: Data diperiksa untuk mengidentifikasi nilai kosong (missing values), data ganda, dan inkonsistensi. Data yang tidak lengkap atau tidak valid ditangani melalui proses penghapusan atau penyesuaian sesuai kebutuhan analisis.
2. Transformasi dan Normalisasi: Karena algoritma SVM sensitif terhadap skala fitur, dilakukan normalisasi menggunakan metode Min-Max Scaling sehingga seluruh variabel berada dalam rentang nilai 0 hingga 1. Normalisasi ini bertujuan untuk menghindari dominasi atribut tertentu yang memiliki skala lebih besar dalam proses pembentukan hyperplane.
3. Seleksi Fitur: Analisis korelasi antarvariabel dilakukan untuk mengidentifikasi atribut yang memiliki hubungan signifikan terhadap tingkat risiko kredit. Fitur yang dianggap relevan dipertahankan untuk digunakan dalam proses pelatihan model.

## 2.3 Pembagian Dataset

Dataset dibagi menjadi dua bagian, yaitu data latih (training set) sebesar 80% dan data uji (testing set) sebesar 20%. Data latih digunakan untuk membangun model klasifikasi SVM, sedangkan data uji digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya. Pembagian ini dilakukan secara proporsional untuk menjaga representasi masing-masing kategori risiko dalam kedua subset data.

## 2.4 Pembangunan Model Support Vector Machine

Model klasifikasi dibangun menggunakan algoritma Support Vector Machine (SVM) dengan pendekatan multi-class classification untuk mengelompokkan nasabah ke dalam tiga kategori risiko. Dalam penelitian ini digunakan tiga jenis kernel yang berbeda untuk membandingkan performa model, yaitu:

1. Kernel Linear, yang membentuk batas keputusan berbasis pemisahan linier.
2. Kernel Polynomial, yang memungkinkan pembentukan batas keputusan berbentuk kurva dengan derajat tertentu.
3. Kernel Radial Basis Function (RBF), yang mampu menangani pola nonlinier dengan memetakan data ke ruang berdimensi lebih tinggi.

Penggunaan beberapa kernel ini bertujuan untuk menganalisis sejauh mana karakteristik data risiko kredit dapat dipisahkan secara linier maupun nonlinier.

## 2.5 Optimalisasi Parameter

Performa model SVM dipengaruhi oleh parameter hiper, khususnya parameter C (cost) dan gamma (untuk kernel nonlinier). Untuk memperoleh konfigurasi parameter yang optimal, dilakukan proses pencarian parameter menggunakan metode Grid Search. Grid Search dilakukan dengan menguji beberapa kombinasi nilai parameter pada data latih untuk menemukan konfigurasi yang menghasilkan performa terbaik. Parameter terbaik yang diperoleh kemudian digunakan dalam proses evaluasi pada data uji.

## 2.6 Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan terhadap data uji menggunakan metrik klasifikasi berikut:

1. Accuracy, untuk mengukur proporsi prediksi yang benar terhadap seluruh data uji.
2. Precision, untuk mengukur ketepatan prediksi pada masing-masing kelas risiko.
3. Recall, untuk mengukur kemampuan model dalam mendeteksi seluruh data pada kelas tertentu.
4. F1-Score, sebagai rata-rata harmonik antara precision dan recall.

Nilai dari masing-masing metrik dibandingkan antar kernel untuk menentukan model dengan performa terbaik dalam mengklasifikasikan tingkat risiko kredit nasabah. Analisis dilakukan dengan menitikberatkan pada keseimbangan precision dan recall, khususnya pada kategori risiko tinggi yang memiliki implikasi signifikan terhadap pengambilan keputusan kredit.

### 3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengimplementasikan algoritma Support Vector Machine (SVM) dalam mengklasifikasikan tingkat risiko kredit nasabah bank berdasarkan data historis dan karakteristik finansial nasabah. Hasil penelitian diperoleh melalui serangkaian tahapan yang telah dijelaskan pada bagian metode, yaitu pengumpulan data, pra-pemrosesan, pembagian dataset, penerapan model SVM dengan beberapa kernel, serta evaluasi performa model menggunakan metrik pengukuran akurasi, presisi, recall, dan F1-score. Bagian ini membahas hasil-hasil yang diperoleh dari penerapan metode tersebut serta interpretasi terhadap performa model yang dihasilkan.

#### 3.1 Hasil Pra Pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk memastikan bahwa data berada dalam kondisi yang layak digunakan untuk pemodelan. Setelah proses pembersihan data dari duplikasi dan nilai kosong, dataset akhir digunakan sebagai dasar pembentukan model klasifikasi. Normalisasi menggunakan metode Min-Max Scaling dilakukan untuk menyamakan rentang nilai seluruh atribut ke dalam interval 0–1. Proses ini penting karena SVM merupakan algoritma yang berbasis perhitungan jarak (distance-based classifier), sehingga perbedaan skala antar fitur dapat memengaruhi pembentukan hyperplane secara signifikan. Tanpa normalisasi, fitur dengan rentang nilai lebih besar berpotensi mendominasi proses optimasi margin. Persamaan rumus untuk normalisasi menggunakan Min-Max Scaling adalah sebagai berikut:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Rumus persamaan (1) digunakan untuk memastikan setiap fitur berada pada skala yang sama, sehingga model SVM tidak berat sebelah terhadap fitur dengan nilai yang besar.

Tabel 1. Distribusi Data dan Kategori Risiko Kredit

Kategori Risiko	Jumlah Data	Persentase
Rendah	220	44%
Sedang	170	34%
Tinggi	110	22%
<b>TOTAL</b>	<b>500</b>	<b>100%</b>

Analisis korelasi antar variabel menunjukkan bahwa atribut seperti pendapatan, jumlah pinjaman, dan rasio utang terhadap pendapatan memiliki kontribusi terhadap pembentukan kategori risiko kredit. Temuan ini konsisten dengan penelitian credit scoring sebelumnya yang menyatakan bahwa variabel kapasitas pembayaran dan beban utang merupakan indikator utama risiko gagal bayar [17].

#### 3.2 Hasil Penerapan Model SVM

Setelah data diproses, model SVM diterapkan menggunakan tiga jenis kernel yaitu Linear, Polynomial, dan Radial Basis Function (RBF). Setiap model diuji menggunakan dataset yang telah dibagi menjadi 80% data latih dan 20% data uji. Parameter utama yang disesuaikan adalah C (cost) dan gamma, yang ditentukan melalui metode grid search agar diperoleh kombinasi optimal. Hasil pengujian menunjukkan bahwa kernel RBF memberikan performa terbaik dibandingkan kernel lainnya. Hal ini karena data klasifikasi risiko kredit bersifat nonlinier dan memiliki distribusi yang kompleks, sehingga kernel RBF lebih mampu membentuk batas pemisah (decision boundary) yang optimal antara kelas risiko rendah, sedang, dan tinggi. Hasil evaluasi performa masing-masing kernel ditampilkan pada tabel 2 berikut ini:

Tabel 2. Perbandingan Hasil Evaluasi Model SVM

No	Jenis Kernel	Akurasi (%)	F1 Score (%)	Presesi (%)	Recall (%)
1	Linear	85.3	82.6	84.1	83.3
2	Polynomial	88.1	86.7	87.2	86.9
3	RFB	91.7	90.5	89.8	90.1

#### 3.3 Proses Pengujian Model

Sebelum dilakukan evaluasi performa, model SVM terlebih dahulu melalui tahapan proses pengujian yang sistematis. Data yang telah dibagi menjadi data latih dan data uji digunakan untuk melatih model dengan berbagai konfigurasi kernel dan parameter. Pada tahap pelatihan, model mempelajari pola hubungan antar variabel untuk membentuk hyperplane pemisah yang optimal. Selanjutnya, model yang telah terbentuk diuji menggunakan data uji

untuk mengukur kemampuan generalisasi terhadap data baru. Proses pengujian ini dilakukan secara berulang untuk setiap jenis kernel (linear, polynomial, dan RBF) dengan kombinasi parameter terbaik hasil optimasi grid search.

Hasil prediksi dari masing-masing model kemudian dibandingkan dengan label aktual untuk membentuk confusion matrix, yang selanjutnya digunakan sebagai dasar perhitungan metrik evaluasi seperti akurasi, precision, recall, dan F1-score. Tahapan ini memastikan bahwa performa model yang dihasilkan tidak hanya optimal pada data latih, tetapi juga stabil dan representatif dalam kondisi nyata.

### 3.4 Analisis Performa Antar Kernel

Berdasarkan Tabel 2, terlihat adanya peningkatan performa dari kernel Linear ke Polynomial, dan mencapai nilai tertinggi pada kernel RBF. Kernel Linear menghasilkan akurasi sebesar 85,3%. Nilai ini menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan benar, namun masih terdapat kesalahan klasifikasi yang relatif tinggi dibandingkan kernel lainnya. Hal ini mengindikasikan bahwa batas keputusan linier belum sepenuhnya mampu memisahkan pola risiko kredit yang bersifat kompleks.

Kernel Polynomial meningkatkan akurasi menjadi 88,1%, atau mengalami peningkatan sebesar 2,8% dibandingkan kernel Linear. Peningkatan ini menunjukkan bahwa pemodelan dengan batas keputusan berbentuk kurva mampu menangkap sebagian pola nonlinier dalam data.

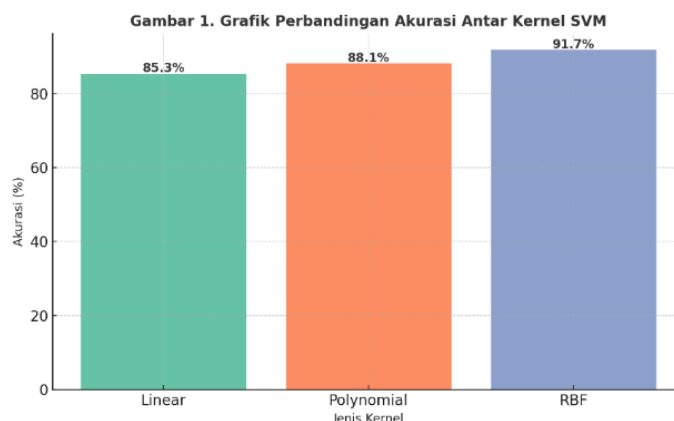
Kernel RBF memberikan performa terbaik dengan akurasi sebesar 91,7%. Jika dibandingkan dengan kernel Linear, terjadi peningkatan sebesar 6,4%, dan dibandingkan dengan Polynomial sebesar 3,6%. Selisih ini menunjukkan bahwa karakteristik data risiko kredit pada perbankan daerah memiliki pola nonlinier yang cukup kuat sehingga lebih optimal dimodelkan menggunakan fungsi kernel berbasis radial.

Selain akurasi, nilai F1-score pada kernel RBF mencapai 90,5%, yang menunjukkan keseimbangan yang baik antara precision dan recall. Hal ini penting dalam klasifikasi risiko kredit, karena model tidak hanya dituntut untuk tepat dalam memprediksi, tetapi juga mampu mendeteksi seluruh kategori risiko secara proporsional. Hasil ini konsisten dengan penelitian sebelumnya yang menyatakan bahwa SVM dengan kernel nonlinier, khususnya RBF, termasuk dalam kelompok model dengan performa terbaik dalam credit scoring dibandingkan pendekatan statistik tradisional [18].

### 3.5. Analisis Precision dan Recall

Precision pada kernel RBF sebesar 89,8% menunjukkan bahwa sebagian besar prediksi yang dihasilkan model sesuai dengan kondisi aktual. Sementara itu, recall sebesar 90,1% mengindikasikan bahwa model mampu mengidentifikasi mayoritas data dalam masing-masing kategori risiko. Dalam lingkup perbankan daerah, nilai recall yang tinggi pada kategori risiko tinggi memiliki implikasi penting karena berkaitan dengan kemampuan bank dalam mendeteksi potensi kredit bermasalah. Model dengan recall rendah berisiko melewatkan nasabah berisiko tinggi, yang dapat meningkatkan rasio kredit macet. Keseimbangan antara precision dan recall yang tercermin dalam nilai F1-score menunjukkan bahwa model tidak bias terhadap salah satu metrik saja, melainkan mempertahankan stabilitas klasifikasi antar kategori risiko.

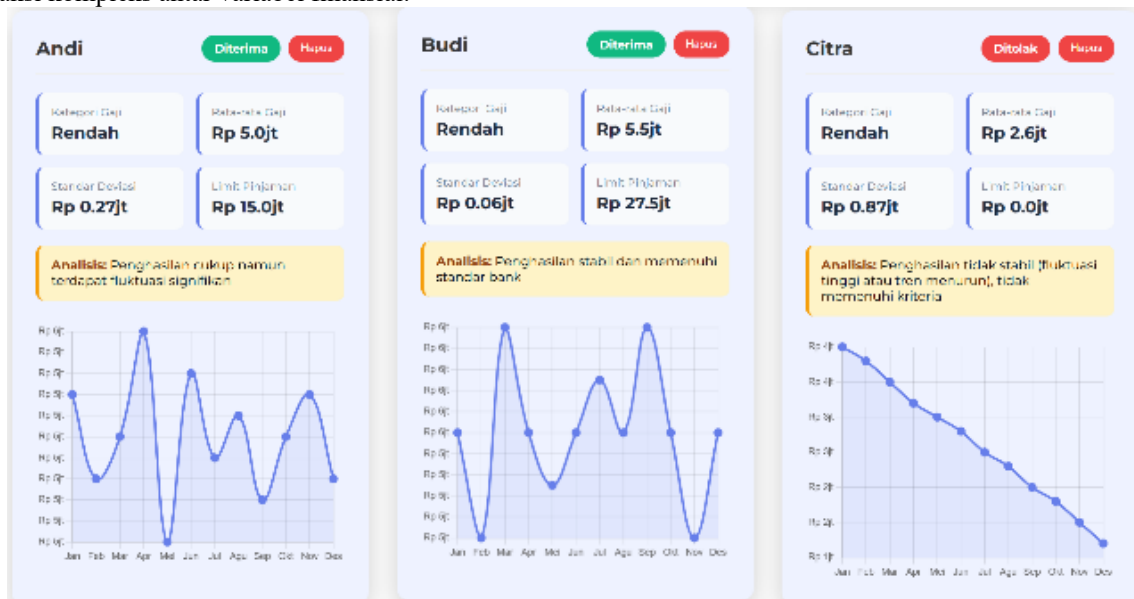
### 3.6. Interpretasi Grafik Perbandingan Kernel



**Gambar 2** Perbandingan Akurasi Antar Kernel

Grafik perbandingan akurasi menunjukkan tren peningkatan performa dari kernel Linear ke Polynomial dan mencapai nilai tertinggi pada kernel RBF. Pola ini konsisten dengan sifat data risiko kredit yang tidak sepenuhnya terpisah secara linier.

Visualisasi hasil kernel RBF menunjukkan bahwa model mampu membentuk batas keputusan yang lebih fleksibel dan adaptif terhadap distribusi data. Hal ini memperkuat temuan bahwa kernel RBF lebih efektif dalam memetakan interaksi kompleks antar variabel finansial.



Gambar 3 Hasil Grafik Menggunakan Kernel RBF

Selain itu, hasil pengujian tidak menunjukkan indikasi overfitting yang signifikan. Model dievaluasi menggunakan data uji yang tidak digunakan dalam proses pelatihan, sehingga nilai akurasi sebesar 91,7% pada kernel RBF merepresentasikan kemampuan generalisasi model terhadap data baru. Jika terjadi overfitting, biasanya model akan menunjukkan performa yang sangat tinggi pada data latih namun menurun secara drastis pada data uji. Dalam penelitian ini, performa antar kernel menunjukkan pola yang stabil dan konsisten, sehingga dapat disimpulkan bahwa peningkatan akurasi pada kernel RBF bukan disebabkan oleh penyesuaian berlebihan terhadap data latih, melainkan karena kemampuan kernel tersebut dalam menangkap struktur nonlinier data.

Perbedaan performa antar kernel juga dapat dianalisis dari selisih nilai akurasi dan F1-score. Kernel RBF menunjukkan peningkatan akurasi sebesar 6,4% dibandingkan kernel linear dan 3,6% dibandingkan kernel polynomial. Dalam konteks klasifikasi risiko kredit, selisih di atas 3% dapat dianggap signifikan secara praktis karena berkaitan langsung dengan jumlah keputusan kredit yang berpotensi berubah. Dengan jumlah data 500 nasabah, peningkatan akurasi sebesar 6,4% setara dengan sekitar 32 prediksi tambahan yang diklasifikasikan secara benar dibandingkan model linear. Secara operasional, perbedaan ini cukup berarti dalam proses evaluasi risiko kredit.

Tren peningkatan performa dari kernel linear ke polynomial dan kemudian ke RBF menunjukkan pola yang konsisten, yaitu semakin fleksibel fungsi kernel dalam memetakan ruang fitur, semakin baik kemampuan model dalam memisahkan kategori risiko. Pola ini mengindikasikan bahwa distribusi data risiko kredit pada perbankan daerah tidak sepenuhnya dapat direpresentasikan oleh model linier sederhana. Interaksi antara variabel seperti pendapatan, rasio utang, dan riwayat pembayaran kemungkinan membentuk batas keputusan yang bersifat kompleks dan tidak simetris, sehingga memerlukan pemetaan nonlinier seperti yang disediakan oleh kernel RBF.

Konsistensi tren ini juga memperkuat argumentasi bahwa pemilihan kernel merupakan faktor krusial dalam implementasi SVM. Jika performa antar kernel tidak menunjukkan pola yang jelas, maka perbedaan nilai akurasi dapat dianggap sebagai fluktuasi acak. Namun dalam penelitian ini, peningkatan performa terjadi secara bertahap dan sistematis, yang menunjukkan bahwa karakteristik data memang lebih sesuai dimodelkan menggunakan pendekatan nonlinier.

Dengan demikian, analisis grafik dan perbandingan kuantitatif antar kernel tidak hanya menunjukkan model terbaik, tetapi juga memberikan pemahaman terhadap struktur data risiko kredit itu sendiri. Temuan ini mendukung literatur sebelumnya yang menyatakan bahwa credit scoring sering kali memiliki batas keputusan nonlinier akibat interaksi variabel finansial yang kompleks.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan Support Vector Machine (SVM) untuk klasifikasi tingkat risiko kredit nasabah bank, dapat disimpulkan bahwa metode SVM terbukti efektif dan andal dalam mengklasifikasikan tingkat risiko kredit berdasarkan data karakteristik nasabah seperti pendapatan, jumlah pinjaman, lama bekerja, riwayat pembayaran, dan rasio utang terhadap pendapatan. Melalui serangkaian tahapan mulai dari pra-

pemrosesan data, pembagian dataset, penerapan model dengan berbagai kernel, hingga evaluasi performa, diperoleh hasil bahwa kernel Radial Basis Function (RBF) menghasilkan performa terbaik dengan tingkat akurasi mencapai 91,7%, presisi 90,5%, dan F1-score 90,1%. Hasil ini menunjukkan bahwa SVM mampu memisahkan kelas risiko kredit dengan margin optimal dan mendeteksi pola nonlinier yang tidak dapat diidentifikasi oleh metode konvensional seperti regresi logistik atau decision tree sederhana. Dengan demikian, SVM dapat dijadikan sebagai pendekatan modern yang efektif untuk mendukung pengambilan keputusan kredit di sektor perbankan, terutama dalam meminimalkan risiko kredit bermasalah (non-performing loan) serta meningkatkan efisiensi penilaian kelayakan nasabah.

Sebagai tindak lanjut dari penelitian ini, disarankan agar pengembangan model dilakukan dengan melibatkan lebih banyak variabel dan data historis jangka panjang, termasuk faktor eksternal seperti kondisi ekonomi makro dan tren inflasi yang dapat memengaruhi kemampuan nasabah dalam membayar pinjaman. Selain itu, penerapan ensemble methods seperti SVM-Boosting atau Hybrid SVM-Neural Network dapat menjadi alternatif untuk meningkatkan akurasi dan stabilitas model. Penelitian selanjutnya juga dapat memperluas analisis menggunakan pendekatan real-time credit scoring system agar hasil prediksi dapat langsung diterapkan pada proses persetujuan kredit di lembaga keuangan. Dengan demikian, pengembangan lanjutan dari model ini diharapkan tidak hanya mampu meningkatkan akurasi klasifikasi, tetapi juga memperkuat sistem manajemen risiko kredit yang lebih adaptif, cerdas, dan berbasis data dalam menghadapi dinamika keuangan modern.

## REFERENCES

- [1] A. Yeriesca, "LAPORAN PROFIL INDUSTRI PERBANKAN - TRIWULAN IV," 2021.
- [2] "Beranda - WEBSITE | PT. BPR SERANG (Perseroda)," BankSerang. Accessed: Jan. 26, 2026. [Online]. Available: <https://www.bankserang.co.id>
- [3] Jasman and E. Susanto, "Manajemen Risiko Kredit Perusahaan Fintech Lending Di Indonesia Berdasarkan Peraturan Nomor 19/SEOJK.06/2023," *Khozana: Jurnal Ekonomi dan Perbankan Islam*, vol. 9, no. 1, pp. 79–89, May 2025, doi: 10.69604/kb5pmd96.
- [4] S. Thakkar, "Credit Risk Assessment With Gradient Boosting Machines," vol. 9, no. 11, 2021.
- [5] R. Hendriadi, A. Putri, D. Amelia, and R. Syafrina, "Design and Development of Credit Scoring Model for Conventional Banks for Individual Borrowing Case Study on PT BPR Sungai Puar District Agam," *JAROE*, vol. 1, no. 1, pp. 43–56, Jul. 2018, doi: 10.24815/jaroe.v1i1.10749.
- [6] G. Zhang, "Comparison of Machine Learning Models and Traditional Models for Forecasting in the Economy," *Computer Life*, vol. 11, no. 3, pp. 1–6, Dec. 2023, doi: 10.54097/n11mbn72.
- [7] E. Dumitrescu, S. Hué, C. Hurlin, and S. Tokpavi, "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects," *European Journal of Operational Research*, vol. 297, no. 3, pp. 1178–1192, Mar. 2022, doi: 10.1016/j.ejor.2021.06.053.
- [8] N. P. D. T. Yanti and I. M. D. P. Asana, "Sistem Klasifikasi Pengajuan Kredit dengan Metode Support Vector Machine (SVM)," *Jurnal Sistem Cerdas*, vol. 6, no. 2, pp. 123–133, Aug. 2023.
- [9] D. K. Silalahi, H. Murfi, and Y. Satria, "Studi Perbandingan Pemilihan Fitur untuk Support Vector Machine pada Klasifikasi Penilaian Risiko Kredit," *EduMatSains : Jurnal Pendidikan, Matematika dan Sains*, vol. 1, no. 2, pp. 119–136, Feb. 2017, doi: 10.33541/edumatsains.v1i2.238.
- [10] B. Pernama and H. D. Purnomo, "Analisis Risiko Pinjaman dengan Metode Support Vector Machine, Artificial Neural Network dan Naïve Bayes," *Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 7, no. 1, pp. 92–99, Jan. 2023, doi: 10.35870/jtik.v7i1.693.
- [11] S. Mestiri, "Credit scoring using machine learning and deep Learning-Based models," *DSFE*, vol. 4, no. 2, pp. 236–248, 2024, doi: 10.3934/DSFE.2024009.
- [12] A. Manik, E. B. Nababan, and T. Tulus, "IMPROVED SUPPORT VECTOR MACHINE PERFORMANCE USING PARTICLE SWARM OPTIMIZATION IN CREDIT RISK CLASSIFICATION," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 6, pp. 1739–1746, Dec. 2022, doi: 10.20884/1.jutif.2022.3.6.615.
- [13] "RBF SVM parameters," scikit-learn. Accessed: Feb. 23, 2026. [Online]. Available: [https://scikit-learn/stable/auto\\_examples/svm/plot\\_rbf\\_parameters.html](https://scikit-learn/stable/auto_examples/svm/plot_rbf_parameters.html)
- [14] A. Rahman, A. Yusnita, and H. Ekawati, "Penerapan Algoritma XGBoost Dalam Prediksi Harga Sewa Kos Di Kota Samarinda," *Bulletin of Information Technology (BIT)*, vol. 6, no. 4, pp. 379–390, Dec. 2025, doi: 10.47065/bit.v7i1.2304.
- [15] F. Muhammad, Suharmanto, J. I. Saputo, and W. S. Utami, "Penerapan Data Mining Menggunakan Metode Cluster K-Means Untuk Pengelompokan Fasilitas Sekolah," *Bulletin of Information Technology (BIT)*, vol. 6, no. 4, pp. 408–416, Dec. 2025, doi: 10.47065/bit.v7i1.2340.
- [16] S. N. Sari, P. Annisa, A. N. D. Rahma, R. P. Ritonga, and D. P. Utomo, "Teknik Data Mining Dengan Menggunakan Algoritma Decision Tree Untuk Mengetahui Pola Pemahaman Mahasiswa Pada Matakuliah



- Pemrograman,” *Bulletin of Information Technology (BIT)*, vol. 6, no. 4, pp. 417–431, Dec. 2025, doi: 10.47065/bit.v6i4.2339.
- [17] W. Komalasari and A. Ardiansyah, “Analisis Risiko Gagal Bayar Piutang Unit Simpan Pinjam Koppeg RSU R. Syamsudin, SH,” *Jurnal Ekualisasi*, vol. 6, no. 2, pp. 47–61, Jul. 2025, doi: 10.60023/p8ae8t34.
- [18] F. Dainelli, G. Bet, and E. Fabrizi, “The financial health of a company and the risk of its default: Back to the future,” *International Review of Financial Analysis*, vol. 95, p. 103449, Oct. 2024, doi: 10.1016/j.irfa.2024.103449.

